



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/610,684	07/02/2003	Scott Shepard	02-4038	2914

7590 02/27/2007  
Leonard C. Suchyta  
c/o Christian Andersen  
Verizon Corporate Services Group Inc.  
600 Hidden Ridge, HQE03H01  
Irving, TX 75038

EXAMINER
----------

SIEDLER, DOROTHY S

ART UNIT	PAPER NUMBER
----------	--------------

2626

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	02/27/2007	PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

<b>Office Action Summary</b>	Application No. 10/610,684	Applicant(s) SHEPARD ET AL.	
	Examiner Dorothy Sarah Siedler	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 02 July 2003.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-46 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-46 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 02 July 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some    \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |   |   |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)  | 5) <input type="checkbox"/> Notice of Informal Patent Application                       |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)<br>Paper No(s)/Mail Date <u>8-25-03, 12-23-03</u> | 6) <input type="checkbox"/> Other: _____  |

### DETAILED ACTION

This is the initial office action in response to the application filled on July 2, 2003.

Claims 1-46 are pending and are considered below.

### ***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims are rejected under 35 U.S.C. 102(e) as being anticipated by ***Saindon*** (6,820,055).

As per claims 1 and 20, ***Saindon*** discloses a method for facilitating translation of an audio signal that includes speech to another language, comprising: retrieving a textual representation of the audio signal (column 1 lines 51-52); presenting the textual representation to a user (column 1 lines 52-53); receiving selection of a segment of the textual representation for translation (column 15 lines 56-55, *the transcriptionist selects a section of the translation for correction*); obtaining a portion of the audio signal corresponding to the segment of the textual representation; providing the segment of the textual representation and the portion of the audio signal to the user (column 15

lines 39-41, *audio synchronized with the text is played back during transcription*); and receiving translation of the portion of the audio signal from the user (column 15 lines 56-55, *the transcriptionist selects a section translation for correction*).

As per claim 21, **Saindon** discloses a translation system, comprising: a memory configured to store instructions (column 5 lines 46-51); and a processor configured to execute the instructions in memory (column 5 lines 41-45) to: obtain a transcription of an audio signal that includes speech (column 1 lines 51-52), present the transcription to a user (column 1 lines 52-53), receive selection of a portion of the transcription for translation (column 15 lines 56-55), retrieve a portion of the audio signal corresponding to the portion of the transcription, provide the portion of the transcription and the portion of the audio signal to the user (column 15 lines 39-41, *audio synchronized with the text is played back during transcription*), and receive translation of the portion of the audio signal from the user (column 15 lines 56-55, *the transcriptionist selects a section translation for correction*).

As per claim 40, **Saindon** discloses a graphical user interface, comprising: a transcription section that includes a transcription of non-text information in a first language (column 1 lines 51-52); a translation section that receives a translation of the non-text information into a second language (column 15 lines 56-55, *the transcriptionist selects a section translation for correction*); and a play button that, when selected,

Art Unit: 2626

causes: retrieval of the non-text information to be initiated, playing of the non-text information, and the playing of the non-text information to be visually synchronized with the transcription in the transcription section (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer. Since the text and audio are played back, it is inherent that there was an initiation method, including a start button*).

As per claims 2 and 22, **Saindon** discloses the method and system of claims 1 and 21, wherein the retrieving a textual representation includes: generating a request for information, sending the request to a server, and obtaining, from the server, at least the textual representation of the audio signal (column 18 lines 35-40, *the processor enables data storage and management of information from a server*).

As per claim 3 and 23, **Saindon** discloses the method and system of claims 1 and 21, wherein the presenting the textual representation to a user, includes: obtaining the audio signal, providing the audio signal and the textual representation of the audio signal to the user, and visually synchronizing the providing of the audio signal with the textual representation of the audio signal (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer. Since the audio is played back, it is inherent that it was first obtained*).

As per claims 5 and 25, **Saindon** discloses the method and system of claims 3 and 23, wherein the obtaining the audio signal includes: receiving input, from the user, regarding a desire for the audio signal (column 1 lines 65-67, *live event audio is received*), initiating a media player, and using the media player to obtain the audio signal (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer*).

As per claims 6 and 26, **Saindon** discloses the method and system of claims 1 and 21, wherein the receiving selection of a segment of the textual representation includes: identifying a portion of the textual representation selected by the user (column 15 lines 56-55, *the transcriptions selects a section translation for correction*), accessing a server to obtain text corresponding to the portion of the textual representation, and receiving, from the server, the text corresponding to the portion of the textual representation (column 18 lines 35-40, *the processor enables data storage and management of information from a server*).

As per claims 8 and 28, **Saindon** discloses the method and system of claims 1 and 21, wherein the obtaining a portion of the audio signal includes: initiating a media player, and using the media player to obtain the portion of the audio signal (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer*).

As per claims 13 and 33, **Saindon** discloses the method and system of claims 1 and 21, wherein the providing the segment of the textual representation and the portion of the audio signal to the user includes: visually synchronizing the providing of the portion of the audio signal with the segment of the textual representation (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer*).

As per claims 14 and 34, **Saindon** discloses the method of claim 13, wherein the segment of the textual representation includes time codes corresponding to when words in the textual representation were spoken (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer. Since the audio and the text are synchronized in time, is inherent that the text has time codes corresponding to the audio data*).

As per claims 19 and 39, **Saindon** discloses the method and system of claims 1 and 21, further comprising: publishing the translation to a user-determined location (column 21 line 66- column 22 line 2).

As per claim 41, **Saindon** discloses the graphical user interface of claim 40 wherein the transcription visually distinguishes names of people, places, and organizations (column 16 lines 34-50, *the transcript is spell checked to determine if proper nouns are capitalized, and corrects them if they are not capitalized*).

As per claim 45, **Saindon** discloses the graphical user interface of claim 40, wherein the non-text information includes at least one of audio and video (column 3 lines 60-67, *the multimedia information, including text and corresponding audio, are played and synchronized through the software viewer*).

### ***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 9, 29, 42, 43, and 46 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Saindon**.

As per claims 9 and 29, **Saindon** discloses the method and system of claims 8 and 28, but does not explicitly state identifying, by the media player, the segment of the textual representation, and retrieving the portion of the audio signal corresponding to the



segment of the textual representation. However, **Saindon** does disclose a system that synchronizes audio and corresponding text information, enabling feedback from a user. In addition, **Saindon** discloses a speech-to-text conversion system where a transcriptionist translates audio information (column 15 lines 56-65).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the media player, synchronizing audio and a corresponding transcription, to determine a segment of text for translation in **Saindon**, since it would provide increased visualization of synchronized data, enabling a quick and efficient selection of text for translation.

As per claim 42, **Saindon** discloses the graphical user interface of claim 40, but does not explicitly disclose a configuration button, that when selected, causes a window to be presented, the window permitting an amount of backup to be specified, the amount of backup including one of a predetermined amount of time and a predetermined number of words. However, **Saindon** does disclose software backup components used by the system (column 17 lines 50-60, *during times of system failure, software or hardware back up is used*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have a configuration button, or a save button, which would enable the backup, or saving, of a specified amounts of information in **Saindon**, since it would

allow the user to periodically back up information, thus avoiding the need to retranslate information due to a sudden system failure, and return to it later to continue the translation or perform edits.

As per claim 43, **Saindon** discloses the graphical user interface of claim 42, but does not explicitly disclose wherein the window further permits a name to be given for the translation and a location of publication to be specified. However, **Saindon** does disclose that a translation can be e-mailed or delivered by a variety of means (column 21 line 66- column 22 line 2).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to give a name to the translation and specify a publication location in **Saindon**, since it would enable the user to save the current progress and return to the translation at a future time to continue the translation or perform edits.

As per claim 46, **Saindon** discloses the graphical user interface of claim 40, but does not explicitly disclose wherein the graphical user interface is associated with a word processing application. However, **Saindon** does disclose a system that synchronizes audio and corresponding text information using a media player integrated with a text viewer, which enables feedback from a user. In addition, **Saindon** discloses a speech-to-text conversion system where a transcriptionist translates audio information (column 15 lines 56-65).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the text viewer integrated into the media player as a word processing application in **Saindon**, since it would provide increased visualization of synchronized data, enabling a quick and efficient selection of text for translation, as well as reducing the number of commands needed to transcribe a translation, since the translation can be carried out in the same window.

Claims 4, 7, 10-12, 15-18, 24, 27, 30-32, 35-38, and 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Saindon** in view of **Schulz** (6,360,237).

As per claims 4 and 24, **Saindon** discloses the method and system of claims 3 and 23, but does not disclose wherein the obtaining the audio signal includes: accessing a database of original media to retrieve the audio signal. **Schulz** discloses accessing a database of original media to retrieve the audio signal (column 4 lines 50-52, *audio recoding*). **Saindon** and **Schulz** both disclose systems for transcription of speech information that synchronize speech and text.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to access a database of original media in **Saindon**, so that system can save speech information, and return to it at a later time to perform a translation.

Art Unit: 2626

As per claims 7 and 27, **Saindon** discloses the method and system of claims 6 and 26, but does not disclose wherein the text includes a transcription of the audio signal and metadata corresponding to the portion of the textual representation. **Schulz** discloses the text including a transcription of the audio signal and metadata corresponding to the portion of the textual representation (column 4 lines 53-59, *a text file created by the speech recognition unit contains the words that were recognized and the beginning and end times for each word*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to include metadata in the transcription of the audio data in **Saindon**, since the metadata can be used to synchronize the text with the audio data, or be used to label pauses which can be removed, as indicated in **Schulz** (column 4 lines 60-65).

As per claims 10 and 30, **Saindon** discloses the method and system of claims 9 and 29, but does not explicitly disclose identifying time codes associated with a beginning and an ending of the segment of the textual representation. However, **Saindon** does disclose the use of time codes for synchronizing audio data with subtitles, i.e. synchronizing audio data with text information (column 4 lines 26-35). The time codes for the audio data are compared with time codes from the translated text then used for synchronization during playback. In addition, **Schulz** discloses a text file including a transcription of the audio signal and metadata corresponding to the portion of the textual representation (column 4 lines 53-59, *a text file created by the speech recognition unit*

*contains the words that were recognized and the beginning and end times for each word).*

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to identify time codes of the segment of text in **Saindon**, in order to synchronize the text with speech during playback, as well as increase the accuracy of translation by identifying a segment to translate without leaving out important words or including partial words.

As per claims 11 and 31, **Saindon** discloses the method and system of claims 9 and 29, but does not explicitly disclose the segment of the textual representation includes a starting position in the textual representation; and wherein the identifying the segment includes: identifying a time code associated with the starting position in the textual representation. However, **Saindon** does disclose the use of time codes for synchronizing audio data with subtitles, i.e. synchronizing audio data with text information (column 4 lines 26-35). The time codes for the audio data are compared with time codes from the translated text then used for synchronization during playback. In addition, **Schulz** discloses the text including a transcription of the audio signal and metadata corresponding to the portion of the textual representation (column 4 lines 53-59, *a text file created by the speech recognition unit contains the words that were recognized and the beginning and end times for each word*).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to identify time codes of the segment of text in **Saindon**, in order to

synchronize the text with speech during playback, as well as increase the accuracy of translation by identifying a segment to translate without leaving out important words or including partial words.

As per claims 12 and 32, **Saindon** discloses the method and system of claims 1 and 21, but does not disclose displaying the segment of the textual representation in a same window as will be used by the user to provide the translation of the portion of the audio signal. **Schulz** discloses a transcription system where the input transcription is displayed in the same window as the target transcription (column 5 lines 28-32, *the text editor that is used to synchronize the audio and the transcript is used to correct errors*). The user performs modification to the text in the same window the text is presented in.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to display the text segment in the same window as will be used to provide the translation in **Saindon**, in order to reduce the amount of commands the user needs to record the translation, since transition between a text window and a translation window is not required, thus increasing the speed of translation.

As per claims 15 and 35, **Saindon** discloses the method and system of claims 14 and 34, but does not explicitly disclose wherein the visually synchronizing the providing of the portion of the audio signal with the segment of the textual representation includes: comparing times corresponding to the providing of the portion of the audio signal to the time codes from the segment of the textual representation, and visually distinguishing

words in the segment of the textual representation when the words are spoken during the providing of the portion of the audio signal. **Schulz** discloses a transcription system where time codes are used to synchronize audio information with text from a transcription, a cursor on the screen used to align the text with the spoken audio being played back (column 6 lines 21-32).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to compare the time codes for the audio signal and textual representation, and visually distinguish words in the text when they are spoken in the audio signal in **Saindon**, since it would provide increased visualization of synchronized data, enabling a quick and efficient selection of text for translation.

As per claims 16,17,18 and 36,37,38, **Saindon** discloses the method of claim 1, but does not explicitly disclose wherein the providing the segment of the textual representation and the portion of the audio signal to the user includes: permitting the user to control the providing of the portion of the audio signal, allowing the user to at least one of fast forward, speed up, slow down, and back up the providing of the portion of the audio signal using foot pedals, and permitting the user to rewind the portion of the audio signal at least one of a predetermined amount of time and a predetermined number of words. **Schulz** discloses a transcription system the permits the user to control the providing of the portion of the audio signal, allowing the user to at least one of fast forward, speed up, slow down, and back up the providing of the portion of the audio signal using foot pedals (column 2 lines 29-32), and permitting the user to rewind

Art Unit: 2626

the portion of the audio signal at least one of a predetermined amount of time and a predetermined number of words (column 2 lines 29-32).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to control the playback of the audio signal using foot pedals in **Saindon**, since it would allow the user to control the playback of the audio, while leaving their hands free to manipulate the text, as indicated in **Schulz** (column 2 lines 29-34).

As per claims 44, **Saindon** discloses the graphical user interface of claim 40, but does not explicitly disclose wherein the play button further causes words in the transcription to be visually distinguished in synchronism with the words in the non-text information being played. **Schulz** discloses a transcription system where time codes are used to synchronize audio information with text from a transcription, a cursor on the screen used to align the text with the spoken audio being played back (column 6 lines 21-32).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the text viewer integrated into the media player as a word processing application in **Saindon**, since it would provide increased visualization of synchronized data, enabling a quick and efficient selection of text for translation, as well as reducing the number of commands needed to transcribe a translation, since the translation can be carried out in the same window.



### ***Conclusion***

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Arase (4,193,119) discloses a system for the translation of foreign language text.
- Shiotani (4,814,988) discloses translation system for translating only part of a specified input region.
- Ellozy (5,649,060) discloses a system for the automatic indexing and aligning of video, audio and text.
- Brown (5,768,603) discloses a system for natural language translation.
- Jachmann (5,146,439) discloses a records management system with transcription capability.

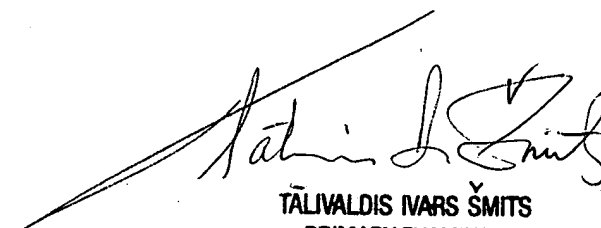
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS



TĀLIVALDIS IVARS ŠMITS  
PRIMARY EXAMINER